Statistics: analyzing & understanding data

Common approach: use parametric model of data

$Bin(p)$, $Poi(\lambda)$, $Exp(\lambda)$, $N(\mu, \sigma^2)$, $Uni[a,b]$

Use $\vec{\theta}$ to denote unknown parameters

Goal: Given indep $\overbrace{\text{samples}}^{x_1, x_2, \ldots x_n}$ from parametric model, determine best choice of parameters $\vec{\theta}$

Approach: Find MLE, most likely choice of $\vec{\theta}$

$$L(x_1, \ldots, x_n | \vec{\theta}) = \prod_{i=1}^{n} f(x_i | \vec{\theta})$$

likelihood function

density cont.
p.m.f. cont.

$$LL(\vec{x} | \vec{\theta}) = \log(L(\vec{x} | \vec{\theta})) = \sum_{i=1}^{n} \log[f(x_i | \vec{\theta})]$$

log-likelihood function

choose $\vec{\theta}$ to maximize $L(\vec{x}|\vec{\theta})$

$\Leftrightarrow$ maximize $LL(\vec{x}|\vec{\theta})$

1 parameter

compute $\frac{dLL}{d\theta}$

set $\frac{dLL}{d\theta} = 0$

Solve

verify soln is max (2^{nd} deriv < 0)

Multiple parameters

$\frac{\partial LL}{\partial \theta_1} = 0$

$\frac{\partial LL}{\partial \theta_2} = 0$

$\vdots$

$\frac{\partial LL}{\partial \theta_k} = 0$

check max

Hessian -ve definite

$\frac{\partial^2 f}{\partial x_i \partial x_j}$

$X_1, \dots, X_n$ samples from $U[0, \theta)$

$\uparrow$
unknown

Find MLE of θ

$$f(x_i) = \begin{cases} \frac{1}{\theta} & 0 \le x_i \le \theta \\ 0 & \text{o.w.} \end{cases}$$

$$L(\vec{x} \mid \theta) = \frac{1}{\theta^n} \quad \text{if all } x_i\text{'s} \in [0, \theta]$$

$$0 \quad \text{otherwise}$$

$$L(\vec{x} \mid \theta) \downarrow \text{ as } \theta \uparrow \quad \Rightarrow \hat{\theta} = \max(x_i)$$

Estimator $\boxed{\text{unbiased}}$ if $E(\hat{\theta}) = \theta$

$\uparrow$
true value

We saw for mean of normal distn

$$E(\hat{\Theta}) = E\left(\frac{\sum X_i}{n}\right) = \mu \qquad \text{unbiased.}$$

for $U[0,\Theta]$ $\quad \hat{\Theta} = \max(X_1, \dots, X_n)$

$E(\max(X_1, \dots, X_n)) = ?$ $\qquad Pr(\hat{\Theta} \le x) = \prod_{i=1}^{n} \underbrace{Pr(X_i \le x)}_{U[0,\Theta]}$

$$= \begin{cases} \frac{x^n}{\Theta^n} & 0 \le x \le \Theta \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\hat{\Theta}}(x) = \begin{cases} \frac{n x^{n-1}}{\Theta^n} & 0 \le x \le \Theta \\ 0 & \text{o.w.} \end{cases}$$

$$E(\hat{\Theta}) = \int_0^{\Theta} x f_{\hat{\Theta}}(x)\,dx = \frac{n}{\Theta^n}\int_0^{\Theta} x \cdot x^{n-1} = \frac{n}{n+1}\Theta$$

$\Rightarrow$ $\hat{\Theta}$ not unbiased, but it is $\boxed{\text{asymptotically unbiased}}$

$$\lim_{n\to\infty} E(\hat{\Theta}_n) = \Theta$$

iid observations $X_1, \dots, X_n$ $U[\theta, \theta+1]$

Find MLE of θ

$$f_{X_i}(x_i) = \begin{cases} 1 & \theta \le x_i \le \theta+1 \\ 0 & \text{o.w.} \end{cases}$$

$$L(\vec{x} \mid \theta) = \begin{cases} 1 & \theta \le \min x_i \le \max x_i \le \theta+1 \\ 0 & \text{otherwise} \end{cases}$$

Any $\hat{\theta} \in [\max X_i - 1, \min X_i]$ equally good

consistent since $\min X_i \longrightarrow \theta$

$\max X_i \longrightarrow \theta+1$

If choose midpoint $\hat{\theta} = \frac{1}{2}[\max X_i + \min X_i - 1]$

it's unbiased

$$E(\hat{\Theta}) = \frac{1}{2}\Big[\underbrace{E(\max X_i)}_{\Theta+1-\frac{1}{n+1}} + \underbrace{E(\min X_i)}_{\Theta+\frac{1}{n+1}} - 1\Big] = \Theta$$

Grade dist'n for 312

$$\begin{cases} A & \frac{1}{2} \\ B & \mu \\ C & 2\mu \\ F & \frac{1}{2}-3\mu \end{cases}$$

MLE for μ when samples are x_A, x_B, x_C, x_F
(Assume we see precise grade each person got)

$$L(x_A, x_B, x_C, x_F \mid \mu) = \left(\frac{1}{2}\right)^{x_A} \mu^{x_B} (2\mu)^{x_C} \left(\frac{1}{2}-3\mu\right)^{x_D}$$

$$\log L(\vec{x} \mid \mu) = x_A \log\left(\tfrac{1}{2}\right) + x_B \log(\mu) + x_C \log(2\mu) + x_D \log\left(\tfrac{1}{2}-3\mu\right)$$

set $\frac{d\,LL}{d\mu} = 0$, solve for $\hat{\mu}$

$$\frac{x_B}{\hat{\mu}} + \frac{2x_C}{2\hat{\mu}} - \frac{3x_D}{\frac{1}{2}-3\hat{\mu}} = 0$$

$$\frac{(x_B + x_C)}{\hat{\mu}} = \frac{3x_D}{\frac{1}{2}-3\hat{\mu}}$$

$$\left(\frac{1}{2}-3\hat{\mu}\right)(x_B+x_C) = 3\hat{\mu}\, x_D$$

$$\frac{x_B+x_C}{2} = 3\hat{\mu}(x_B+x_C+x_D)$$

$$\hat{\mu} = \frac{x_B+x_C}{6(x_B+x_C+x_D)}$$

S